Solutions to review questions #2

I use parentheses (as in $\begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix}$) instead of Strang's brackets (as in $\begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}$) for matrices and vectors. As a consequence, when I write (a, b, c), I mean the row vector $\begin{bmatrix} a & b & c \end{bmatrix}$, and not (an abbreviation for) the column vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Sorry for this! I am just more used to parentheses, and if I try changing my notations, chances are you'll see a mix of both of them in the below.

(When I want to write the column vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ in a compact form, I write $(a, b, c)^T$.)

ad problem 1: (a) Let $v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ and $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$ be our three vectors. Gram-Schmidt orthogonalization will yield three vectors q_1, q_2, q_3 given by the formulas

$$\begin{aligned} q_1 &= v_1; \\ q_2 &= v_2 - \frac{q_1^T v_2}{q_1^T q_1} q_1; \\ q_3 &= v_3 - \frac{q_1^T v_3}{q_1^T q_1} q_1 - \frac{q_2^T v_3}{q_2^T q_2} q_2. \end{aligned}$$

(These are the same formulas as the equality $A = a$ and the equalities (7) and (8) given on page 234, but here we call $v_1, v_2, v_3, q_1, q_2, q_3$ what has been called

a, b, c, A, B, C in the book.[1]) Plugging in, we obtain

$$q_1 = v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix};$$

$$q_2 = v_2 - \frac{q_1^T v_2}{q_1^T q_1} q_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \\ 0 \end{pmatrix};$$

$$q_3 = v_3 - \frac{q_1^T v_3}{q_1^T q_1} q_1 - \frac{q_2^T v_3}{q_2^T q_2} q_2 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ -1 \end{pmatrix}.$$

These three vectors are only orthogonal so far, not yet orthonormal. To make them orthonormal, divide each of them by its length, thus obtaining $\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} \frac{1}{6}\sqrt{6} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ 0 \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{6}\sqrt{3} \\ \frac{1}{6}\sqrt{3} \\ \frac{1}{6}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} \end{pmatrix}$.

[**Remark:** There is a pattern here. Applying Gram-Schmidt orthogonaliza-

[1]If we had m vectors $v_1, v_2, ..., v_m$ instead of three vectors v_1, v_2, v_3, then the corresponding m equations for the $q_1, q_2, ..., q_m$ would look like this:

$$q_i = v_i - \frac{q_1^T v_i}{q_1^T q_1} q_1 - \frac{q_2^T v_i}{q_2^T q_2} q_2 - \cdots - \frac{q_{i-1}^T v_i}{q_{i-1}^T q_{i-1}} q_{i-1}.$$

tion (without normalizing the vectors to length 1) to the $n-1$ vectors $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,

$\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, ..., $\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}$ in $\mathbb{R}^n$, we obtain the $n-1$ vectors $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,

..., $\begin{pmatrix} \frac{1}{n-1} \\ \frac{1}{n-1} \\ \frac{1}{n-1} \\ \vdots \\ \frac{1}{n-1} \\ -1 \end{pmatrix}$ (the i-th vector consists of i coordinates equal to $\frac{1}{i}$, then a single coordinate equal to -1, and all remaining coordinates are 0). This can be proven by checking that these $n-1$ vectors are mutually orthogonal and the i-th of them is a linear combination of the first i of the original vectors.]

(b) $A = \begin{pmatrix} 15 & 6 \\ 8 & 61 \end{pmatrix} = QR$ for $Q = \begin{pmatrix} \frac{15}{17} & -\frac{8}{17} \\ \frac{8}{17} & \frac{15}{17} \end{pmatrix}$ and $R = \begin{pmatrix} 17 & 34 \\ 0 & 51 \end{pmatrix}$.

To find this, apply Gram-Schmidt orthogonalization (including normalizing the lengths to 1) to the two columns of A (the result is $q_1 = \begin{pmatrix} \frac{15}{17} \\ \frac{8}{17} \end{pmatrix}$ and $q_2 = \begin{pmatrix} -\frac{8}{17} \\ \frac{15}{17} \end{pmatrix}$), and then use the formula (9) on page 236 (but here, the matrices are 2×2).

The inverses are $Q^{-1} = \begin{pmatrix} \frac{15}{17} & \frac{8}{17} \\ -\frac{8}{17} & \frac{15}{17} \end{pmatrix}$, $R^{-1} = \begin{pmatrix} \frac{1}{17} & -\frac{2}{51} \\ 0 & \frac{1}{51} \end{pmatrix}$ and $A^{-1} =$

$\begin{pmatrix} \frac{61}{867} & -\frac{2}{289} \\ -\frac{8}{867} & \frac{5}{289} \end{pmatrix}$. An easy way to invert Q is to recall that Q is orthogonal, whence $Q^{-1} = Q^T$. An easy way to invert A is to recall that $A = QR \Longrightarrow A^{-1} = (QR)^{-1} = R^{-1}Q^{-1}$.

ad problem 2: (a) The four $*$'s stand for unknowns (not necessarily equal); let us call them x, y, z, w (from top to bottom) instead. The matrix Q then becomes

$$Q = c \begin{pmatrix} 1 & -1 & -1 & x \\ -1 & 1 & -1 & y \\ -1 & -1 & -1 & z \\ -1 & -1 & 1 & w \end{pmatrix}.$$

For Q to be orthogonal, the columns of Q have to be orthogonal. Equivalently, the columns of the matrix $\begin{pmatrix} 1 & -1 & -1 & x \\ -1 & 1 & -1 & y \\ -1 & -1 & -1 & z \\ -1 & -1 & 1 & w \end{pmatrix}$ have to be orthogonal (because the scaling factor c does not matter, unless it is 0 in which case Q surely will not be orthogonal). It is easy to see that the first three columns of this matrix already are orthogonal, so it only remains to choose x, y, z, w such that the fourth column is orthogonal to them all. In other words, we must have

$$\begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0; \qquad \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0; \qquad \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix}^T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0.$$

This rewrites as a system of linear equations for x, y, z, w:

$$\begin{cases} x - y - z - w = 0; \\ -x + y - z - w = 0; \\ -x - y - z + w = 0 \end{cases}.$$

The solutions of this system are scalar multiples of the vector $\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$, so we can definitely set $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$ for some scalar $\lambda \in \mathbb{R}$. Thus, $x = \lambda 1 = \lambda$, $y = \lambda 1 = \lambda$, $z = \lambda (-1) = -\lambda$ and $w = \lambda 1 = \lambda$.

Thus,

$$Q = c \begin{pmatrix} 1 & -1 & -1 & x \\ -1 & 1 & -1 & y \\ -1 & -1 & -1 & z \\ -1 & -1 & 1 & w \end{pmatrix} = c \begin{pmatrix} 1 & -1 & -1 & \lambda \\ -1 & 1 & -1 & \lambda \\ -1 & -1 & -1 & -\lambda \\ -1 & -1 & 1 & \lambda \end{pmatrix}$$

$$= \begin{pmatrix} c & -c & -c & \lambda c \\ -c & c & -c & \lambda c \\ -c & -c & -c & -\lambda c \\ -c & -c & c & \lambda c \end{pmatrix}. \qquad (1)$$

Now, for Q to be orthogonal, not only must the columns of Q be mutually orthogonal; they also have to have length 1. But the lengths of the four columns of the matrix on the right hand side of (1) are $2|c|$, $2|c|$, $2|c|$ and $2|\lambda c|$ (for example, $\left\| \begin{pmatrix} c \\ -c \\ -c \\ -c \end{pmatrix} \right\| = \sqrt{c^2 + (-c)^2 + (-c)^2 + (-c)^2} = \sqrt{4c^2} = 2\sqrt{c^2} = 2|c|$; do not forget the absolute values!). So $2|c|$, $2|c|$, $2|c|$ and $2|\lambda c|$ must be 1. In other words, $|c| = |\lambda c| = \frac{1}{2}$. This gives rise to four solutions:

$$\left(c = \frac{1}{2} \text{ and } \lambda = 1 \right); \qquad \left(c = \frac{1}{2} \text{ and } \lambda = -1 \right);$$
$$\left(c = -\frac{1}{2} \text{ and } \lambda = 1 \right); \qquad \left(c = -\frac{1}{2} \text{ and } \lambda = -1 \right).$$

These result in the following four values of Q:

$$Q = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix}; \qquad Q = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 \end{pmatrix};$$
$$Q = -\frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix}; \qquad Q = -\frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 \end{pmatrix}.$$

It is easy to check that all these four matrices are indeed orthogonal (and distinct).

(Does the exercise ask for all of them or one of them? I don't know.)

(b) This does not depend on which of the four possible choices for Q we take, because these choices only differ in the signs of the columns (and these don't matter when projecting). Let us take the first choice:

$$Q = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix}.$$

Then, the first column of Q is $\frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$. Since this column is orthonormal, we can compute the projection of b onto this column using formula (5) on page 233 (with $n = 1$ and $q_1 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$). We obtain

$$p = q_1 \left(q_1^T b \right) = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

as the projection of b onto the first column of Q.

To project b onto the plane spanned by the first two columns of Q, we use formula (5) on page 233 again, taking $n = 2$ and taking q_1 and q_2 to be the first two columns $\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ and $\begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ of Q. The resulting projection is

$$p = q_1 \left(q_1^T b \right) + q_2 \left(q_2^T b \right) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

(c) So we are to run the Gram-Schmidt algorithm on the columns

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

of the matrix $A = \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix}$. Let us first make the columns orthogonal, and only then normalize them to have length 1. The first three columns v_1, v_2, v_3 of A are already mutually orthogonal, so they survive the orthogonalization unchanged:

$$q_1 = v_1, \qquad q_2 = v_2, \qquad q_3 = v_3.$$

The fourth column gives rise to the fourth vector

$$q_4 = v_4 - \frac{q_1^T v_4}{q_1^T q_1} q_1 - \frac{q_2^T v_4}{q_2^T q_2} q_2 - \frac{q_3^T v_4}{q_3^T q_3} q_3 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

These four vectors q_1, q_2, q_3, q_4 are mutually orthogonal, but not orthonormal. To get orthonormal vectors, we have to divide them by their lengths:

$$\frac{q_1}{||q_1||} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \qquad \frac{q_2}{||q_2||} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \qquad \frac{q_3}{||q_3||} = \frac{1}{2} \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

$$\text{and} \qquad \frac{q_4}{||q_4||} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

[Notice that the lengths of all four vectors were 2. This was a happy coincidence; most often these lengths will be distinct and contain square roots.]

ad problem 3: (a) We can write the system as $Ax = b$ for $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. To find the least-squares solution, we follow the strategy on page 218 and solve $A^T A\widehat{x} = A^T b$ for $\widehat{x}$. This rewrites as $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \widehat{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and the solution is $\widehat{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

(b) The equations are

$$\begin{aligned} 7 &= C + D(-1); \\ 7 &= C + D1; \\ 21 &= C + D2. \end{aligned}$$

In other words, $\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 21 \end{pmatrix}$. In yet other words, $Ax = b$ for $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$, $x = \begin{pmatrix} C \\ D \end{pmatrix}$ and $b = \begin{pmatrix} 7 \\ 7 \\ 21 \end{pmatrix}$. To find the least-squares solution, we follow the strategy on page 218 and solve $A^T A\widehat{x} = A^T b$ for $\widehat{x}$. This rewrites as $\begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \widehat{x} = \begin{pmatrix} 35 \\ 42 \end{pmatrix}$, and the solution is $\widehat{x} = \begin{pmatrix} 9 \\ 4 \end{pmatrix}$. That is, $C = 9$ and $D = 4$.

ad problem 4: (a) Using the big formula:

$$\det A = 1 \cdot 3 \cdot 1 + 1 \cdot 1 \cdot 4 + 2 \cdot 1 \cdot 1 - 1 \cdot 1 \cdot 1 - 1 \cdot 1 \cdot 1 - 2 \cdot 3 \cdot 4 = -17.$$

(b) Here it is easier to first simplify the determinant:

$$\det B = \det \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 4 & 1 & 1 \\ 5 & 1 & 1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 4 & 1 & 1 \\ 4 & -3 & 0 & 0 \end{pmatrix}$$

(here, we subtracted row 3 from row 4)

$$= \det \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 0 & 3 & -2 & 0 \\ 4 & -3 & 0 & 0 \end{pmatrix} \qquad \text{(here, we subtracted row 2 from row 3)}$$

$$= \det \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 2 & -1 \\ 0 & 3 & -2 & 0 \\ 4 & -3 & 0 & 0 \end{pmatrix} \qquad \text{(here, we subtracted row 1 from row 2)}$$

$$= \underbrace{(-1)^{4+1}}_{=-1} \underbrace{4 \det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & -1 \\ 3 & -2 & 0 \end{pmatrix}}_{=-17} + \underbrace{(-1)^{4+2}}_{=1} (-3) \underbrace{\det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & -2 & 0 \end{pmatrix}}_{=-2}$$

$$+ \underbrace{(-1)^{4+3} 0 \det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 3 & 0 \end{pmatrix}}_{=0} + \underbrace{(-1)^{4+4} 0 \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 3 & -2 \end{pmatrix}}_{=0}$$

(by cofactor expansion in the fourth row)

$$= 74.$$

[**Remark:** There is a pattern to these determinants. The determinant of the $n \times n$-matrix $\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 2 \\ 1 & 1 & 1 & \cdots & 3 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & n-1 & \cdots & 1 & 1 \\ 1 & n & 1 & \cdots & 1 & 1 \\ n+1 & 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$ is $(-1)^{\lfloor n/2 \rfloor} n! \left(1 + \sum\limits_{k=1}^{n} \frac{1}{k}\right)$. This can be proven by induction over n; the idea is to clear out most of the 1's from the matrix by subtracting row 2 from row 1, row 3 from row 2, etc., row n from row $n-1$, and then applying the cofactor expansion with respect to the first column, and recalling that the determinant of a triangular matrix is the product of its diagonal entries.]

(c) The matrix B is invertible since its determinant $\det B = 74$ is nonzero.

To find the entry $(1,4)$ of the inverse, use formula (6) on page 270. It gives

$\left(B^{-1}\right)_{1,4} = \frac{C_{4,1}}{\det B}$, where

$$C_{4,1} = \underbrace{(-1)^{4+1}}_{=-1} \underbrace{\det \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 1 \\ 4 & 1 & 1 \end{pmatrix}}_{=-17} = 17.$$

Thus (and because of $\det B = 74$), we have

$$\left(B^{-1}\right)_{1,4} = \frac{C_{4,1}}{\det B} = \frac{17}{74}.$$

ad problem 5: See http://web.mit.edu/18.06/www/Fall09/exam2soln.pdf.

ad problem 6: (a) One way to do is using the big formula, which (for an arbitrary $n \times n$-matrix $\left(a_{i,j}\right)_{1 \leq i,j \leq n}$) looks as follows:

$$\det\left(\left(a_{i,j}\right)_{1 \leq i,j \leq n}\right) = \sum_{\pi \text{ is a permutation of } \{1,2,\ldots,n\}} (-1)^{\pi} a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}.$$

Each addend of this sum corresponds to a way to pick an entry of row 1, an entry of row 2, etc., an entry of row n, such that no two entries lie in the same column.[2]

For our peculiar matrix $\begin{pmatrix} a & b & c & d \\ l & 0 & 0 & e \\ k & 0 & 0 & f \\ j & i & h & g \end{pmatrix}$, most of these addends are 0. We can, of course, restrict ourselves to the nonzero addends. To obtain a nonzero addend, one has to pick an entry of row 1, an entry of row 2, etc., an entry of row n, such that no two entries lie in the same column, and such that no 0 entry is picked. This doesn't leave us many choices: in fact, we must pick either l or e from row 2, which then forces us to pick f (if we have picked l) or k (if we have picked e) from row 3 (because we must not pick two entries lying in the same column); then, our only choices in row 1 are b and c, and correspondingly we are forced to pick h (if we took b) or i (if we took c) from row 4. Altogether, we get four addends:

- an addend $blfh$ corresponding to the permutation $(2,1,4,3)$ (because it picks the 2nd entry of row 1, the 1st entry of row 2, etc.) with sign 1;

[2] The notation $(-1)^{\pi}$ stands for the *sign* of the permutation π; it is 1 if the list $(\pi(1), \pi(2), \ldots, \pi(n))$ is obtained from the list $(1,2,\ldots,n)$ by an even number of switches, and -1 if it is obtained by an odd number of switches. This sign is the $\det P$ in formula (8) on page 258 of the book. Other notations for this sign are $\operatorname{sign} \pi$ and $\operatorname{sgn} \pi$. I do not know which of these notations was used in class.

- an addend $-clfi$ corresponding to the permutation $(3,1,4,2)$ (because it picks the 3rd entry of row 1, etc.) with sign -1;
- an addend $-bekh$ corresponding to the permutation $(2,4,1,3)$ with sign -1;
- an addend $ceki$ corresponding to the permutation $(3,4,1,2)$ with sign 1.

The big formula thus shows that the determinant is

$$blfh - clfi - bekh + ceki = (bh - ci)(lf - ek).$$

[You are not required to find the factorization.]

Here is an *alternative solution:* Recall that the determinant of a matrix changes sign every time we switch two rows or switch two columns. Thus,

$$\begin{aligned}
\det \begin{pmatrix} a & b & c & d \\ l & 0 & 0 & e \\ k & 0 & 0 & f \\ j & i & h & g \end{pmatrix} &= -\det \begin{pmatrix} a & b & c & d \\ j & i & h & g \\ k & 0 & 0 & f \\ l & 0 & 0 & e \end{pmatrix} && \text{(here, we switched two rows)} \\
&= \det \begin{pmatrix} c & b & a & d \\ h & i & j & g \\ 0 & 0 & k & f \\ 0 & 0 & l & e \end{pmatrix} && \text{(here, we switched two columns)} \\
&= \det \begin{pmatrix} c & b \\ h & i \end{pmatrix} \det \begin{pmatrix} k & f \\ l & e \end{pmatrix} && \text{(by Exercise 23 (a) on page 266)} \\
&= (ci - bh)(ke - fl) = (bh - ci)(lf - ek).
\end{aligned}$$

This gives the result in the factorized form.

(b) Use the big formula as in part **(a)**, but notice that the analysis becomes simpler (despite the matrix being larger!): there is no way to pick nonzero entries from each row in such a way that no two entries are picked from the same column. In fact, we can only pick p or f from row 2; then, we can only pick g (if we chose p) or o (if we chose p) from row 3; then, there are nonzero entries left to pick from row 4 anymore. Hence, there are no nonzero terms in the big formula. Consequently, the determinant is 0.

ad problem 7: Let us simplify the determinant by subtracting row 1 from row 3 and subtracting row 2 from row 4 (we do these two operations in one step because they don't interfere with each other):

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 8 & 7 & 6 & 5 \\ 9 & 10 & 11 & 12 \\ 16 & 15 & 14 & 13 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 8 & 7 & 6 & 5 \\ 8 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 \end{pmatrix}.$$

The matrix on the right hand side has two equal rows, and so its determinant is 0 (because equal rows are linearly dependent). Consequently, the determinant we are looking for is 0.

[**Remark:** The same argument goes through for the $n \times n$-matrix which is filled in in the same way as our 4×4-matrix, for every $n \geq 4$. The determinant is 0.]

ad problem 8: (a) If P is the projection matrix onto a subspace V, then V is the column space of P [3] The subspace we are looking for is therefore the column space of our 4×4-matrix. To find its basis, it is enough to find the rank r of our matrix and pick r linearly independent columns of the matrix. This is straightforward: $r = 2$, and we can take (for example) the first two columns of the projection matrix to obtain a basis. (Actually, any two columns would work, except for the second and third column.)

(b) The formula for the projection matrix P on the line spanned by a column vector a is $P = \frac{aa^T}{a^T a}$ (see the very bottom of page 208). Setting $a = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$, we obtain $P = \begin{pmatrix} \frac{1}{6} & \frac{2}{6} & \frac{-1}{6} \\ \frac{2}{6} & \frac{4}{6} & \frac{-2}{6} \\ \frac{-1}{6} & \frac{-2}{6} & \frac{1}{6} \end{pmatrix}$.

(c) Of course, one can do this the usual way (finding the characteristic polynomial of P, then the eigenvalues, then the eigenvectors), but it is much easier to remember that P is a projection matrix on the line L spanned by the column vector $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. Thus, every vector on the line L (in particular, the vector $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ itself) is an eigenvector for eigenvalue 1 (because the projection matrix P sends this vector to itself), whereas every vector orthogonal to the line L is an eigenvector for eigenvalue 0 (because the projection matrix P sends it to 0). In more detail:

- The vectors in L are eigenvectors of P for eigenvalue 1; this gives us one linearly independent eigenvector $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.

[3] Why?

By the definition of the projection matrix, Pv is the projection of v on V for every $v \in \mathbb{R}^n$. As a consequence, $Pv \in V$ for every $v \in \mathbb{R}^n$. This shows that the column space of P is contained in V. Conversely, every vector $v \in V$ lies in the column space of P (because for every vector v, the projection of v on V is v itself, and so we have $v = Pv$), and so the column space of P contains V. Thus, V and the column space of P are identical, qed.

- The vectors orthogonal to L (that is, the vectors in $L^\perp$) are eigenvectors of P for eigenvalue 0; this gives us two linearly independent eigenvectors $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. (These two vectors are a basis of $L^\perp$; any other basis would do the trick just as well.[4])

Altogether, we have thus obtained three eigenvectors $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ for the eigenvalues 1, 0 and 0, respectively. These three eigenvectors are easily seen to be linearly independent, and so we can diagonalize P using the formula (1) on page 298. So let S be the eigenvector matrix $\begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ whose columns are these three eigenvectors is invertible. Then, $S^{-1}PS$ is the eigenvalue matrix $\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Now, we can diagonalize P as follows: $P = S\Lambda S^{-1}$.

(Of course, other choices of S and Λ are also possible, though the diagonal entries of Λ will always be $1, 0, 0$ in some order.)

(d) See http://web.mit.edu/18.06/www/Spring05/exam2sol.pdf.

ad problem 9:

(a) The eigenvalues of A are the roots of the characteristic polynomial of A, which is $\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{pmatrix} = (1-\lambda)(3-\lambda) - 4 \cdot 2 = \lambda^2 - 4\lambda - 5 = (\lambda+1)(\lambda-5)$ (where λ is the indeterminate). Thus, the eigenvalues of A are -1 and 5. The respective eigenvectors are:

- all multiples of $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ for eigenvalue -1;
- all multiples of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for eigenvalue 5.

(To obtain this, it suffices to recall that the eigenvectors for eigenvalue -1 are the vectors in the nullspace of $A - (-1)I = A + I$, wheras the eigenvectors for eigenvalue 5 are the vectors in the nullspace of $A - 5I$. You know how to find a basis of a nullspace.)

[4] A basis of $L^\perp$ can be found using Gaussian elimination, as $L^\perp$ is the nullspace of the 1×3-matrix $\begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$.

The matrix $A+I$ has the <u>same</u> eigenvectors as A. Its eigenvalues are <u>greater by 1</u>. (This is because if $Av = \lambda v$, then $(A+I)\,v = Av + v = \lambda v + v = (\lambda+1)\,v$.)

(b) This is solved in the same way as problem 8 (c) (with a new matrix), so I will just give one answer (again, there are multiple possible answers because one has freedom in choosing the eigenvectors):

- for the eigenvalue 1, two linearly independent eigenvectors are $\begin{pmatrix}1\\2\\0\end{pmatrix}$ and $\begin{pmatrix}0\\0\\1\end{pmatrix}$;
- for the eigenvalue 0, one linearly independent eigenvector is $\begin{pmatrix}1\\-\frac{1}{2}\\0\end{pmatrix}$.

ad problem 10: See http://web.mit.edu/18.06/www/Fall04/q2sol.pdf.

ad problem 11: (a) See page 537.

(b) If $A = S\Lambda S^{-1}$, then $A^3 = S\Lambda^3 S^{-1}$ and $A^{-1} = S\Lambda^{-1}S^{-1}$.

(Generally, $A^k = S\Lambda^k S^{-1}$ for every integer k. When k is positive, this follows from the computation $A^k = \left(S\Lambda S^{-1}\right)^k = \left(S\Lambda S^{-1}\right)\left(S\Lambda S^{-1}\right)\cdots\left(S\Lambda S^{-1}\right) = S\Lambda \underbrace{\left(S^{-1}S\right)}_{=I} \Lambda \underbrace{\left(S^{-1}S\right)}_{=I} \cdots \underbrace{\left(S^{-1}S\right)}_{=I} \Lambda S^{-1} = S\Lambda\Lambda\cdots\Lambda S^{-1} = S\Lambda^k S^{-1}$. For negative k, it can be proven by showing that $S\Lambda^k S^{-1}$ and $S\Lambda^{-k}S^{-1}$ are mutually inverse.)

(c) The diagonalization of A is $A = S\Lambda S^{-1}$ for $S = \begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}$ and $\Lambda = \begin{pmatrix}3 & 0\\0 & 1\end{pmatrix}$. (Again, this is only one of many possible correct answers.) Since Λ is a diagonal matrix, its powers are easily computed: $\Lambda^k = \begin{pmatrix}3^k & 0\\0 & 1^k\end{pmatrix} = \begin{pmatrix}3^k & 0\\0 & 1\end{pmatrix}$. Now, the powers of A are obtained by the formula $A^k = S\Lambda^k S^{-1}$ (that we have seen above in the solution to part **(a)**):

$$A^k = S\Lambda^k S^{-1} = \begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}\begin{pmatrix}3^k & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix}1+3^k & 1-3^k\\1-3^k & 1+3^k\end{pmatrix}$$

(after straightforward computation).

ad problem 12:

(a) Let $A = \begin{pmatrix} -9 & 8 \\ -10 & 9 \end{pmatrix}$. Then, we are looking for A^{20}. We can diagonalize A as $A = S\Lambda S^{-1}$ with $S = \begin{pmatrix} 4 & 1 \\ 5 & 1 \end{pmatrix}$ and $\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Since Λ is a diagonal matrix, its powers are easily computed: $\Lambda^k = \begin{pmatrix} 1^k & 0 \\ 0 & (-1)^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & (-1)^k \end{pmatrix}$. Now, the powers of A can be obtained by the formula $A^k = S\Lambda^k S^{-1}$ (that we have seen above in the solution to problem 11 **(a)**):

$$A^k = S\Lambda^k S^{-1} = \begin{pmatrix} 4 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 1^k & 0 \\ 0 & (-1)^k \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 5 & 1 \end{pmatrix}^{-1}.$$

You can simplify this, but since you are only looking for A^{20}, there is a simpler way: notice that $\Lambda^{20} = \begin{pmatrix} 1 & 0 \\ 0 & (-1)^{20} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$, so that $A^{20} = S \underbrace{\Lambda^{20}}_{=I_2} S^{-1} = SI_2S^{-1} = SS^{-1} = I_2$.

(An alternative solution proceeds by realizing that $A^2 = I_2$ (by a simple computation), and thus $A^{20} = \left(A^2\right)^{10} = I_2^{10} = I_2$. But how would you guess that A^2 is something as simple as I_2 ? Diagonalization makes it obvious.)

(b) Let $u_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $u_\infty = \lim\limits_{k\to\infty} A^k u_0$. Then, u_∞ is the vector that we are looking for.

The matrix A is Markov. While A is not positive (so the blue box on page 433 does not apply verbatim), A still has the property that it has the eigenvalue 1 with (algebraic) multiplicity 1. Hence, up to scalar multiplication, there is a unique eigenvector of A for eigenvalue 1. This eigenvector is $v = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{2} \end{pmatrix}$ (or any scalar multiple of it).

It might appear that I have pulled the eigenvector – as well as its uniqueness – out of my hat. But it follows from straightforward computations: **Every** Markov matrix has 1 as its eigenvalue at least once. To find all eigenvalues of A for 1, we need to find the nullspace of $A - 1I = A - I$. This can be done by Gaussian elimination (it is a system of 5 equations in 5 indeterminates), and the result is that the nullspace of $A - I$ is 1-dimensional, with basis consisting of the vector

$\begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{2} \end{pmatrix}$ (or any nonzero multiple of it). That's how I got the eigenvector.[5]

Now, $u_\infty = \lim\limits_{k\to\infty} A^k u_0$, so that $Au_\infty = A \lim\limits_{k\to\infty} A^k u_0 = \lim\limits_{k\to\infty} AA^k u_0 = \lim\limits_{k\to\infty} A^{k+1} u_0 = \lim\limits_{k\to\infty} A^k u_0 = u_\infty$. Thus, u_∞ is an eigenvector of A for eigenvalue 1, therefore a scalar multiple of v. In other words, $u_\infty = cv$ for some $c \in \mathbb{R}$. It remains to find the scalar c.

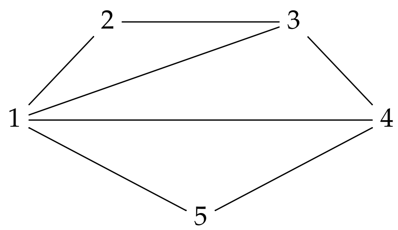
Let j be the row vector $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$. Then, $jA = j$ (since every column of A has sum 1); thus, $j = jA = jAA = jA^2 = jAA^2 = jA^3 = \cdots$. That is, $j = jA^k$ for

---

[5]There is a simpler way to find the eigenvector if you know the following theorem from class:

> **Theorem 0.1.** Let G be a connected (undirected) graph whose vertices are labelled $1, 2, \ldots, n$. For every $i \in \{1, 2, \ldots, n\}$, let $u_i \in \mathbb{R}^n$ be the column vector whose k-th entry (for each $k \in \{1, 2, \ldots, n\}$) is
>
> $$\begin{cases} \dfrac{1}{\text{the number of edges of } G \text{ from } i}, & \text{if } k \text{ is connected to } i; \\ 0, & \text{otherwise} \end{cases}.$$
>
> Let B be the $n \times n$ matrix whose columns are $u_1, u_2, \ldots, u_n$. Then, the matrix B has a unique (up to scaling) eigenvector for the eigenvalue 1. This eigenvector is the vector whose k-th coordinate (for each $k \in \{1, 2, \ldots, n\}$) is the number of edges of G from k.

Our matrix A is precisely the matrix B obtained from the graph $G =$



. So Theorem 0.1 says that this matrix A has a unique (up to scaling) eigenvector for the eigenvalue 1, and this eigenvector is the vector whose k-th coordinate (for each $k \in \{1, 2, \ldots, n\}$) is the number of edges of G from k. This eigenvector is

$$\begin{pmatrix} 4 \\ 2 \\ 3 \\ 3 \\ 2 \end{pmatrix}.$$

This is exactly the same vector as the v that we are found, up to a scalar factor of 4 (and eigenvectors always can be modified by scalar factors).

every $k \in \mathbb{N}$. Hence, $ju_0 = \lim_{k\to\infty} \underbrace{ju_0}_{=jA^k u_0} = \lim_{k\to\infty} jA^k u_0 = j \underbrace{\lim_{k\to\infty} A^k u_0}_{=u_\infty = cv} = jcv = cjv$.

Since $ju_0 = 1$ and $jv = \dfrac{7}{2}$, this becomes $1 = c \cdot \dfrac{7}{2}$, so that $c = \dfrac{2}{7}$ and thus

$$u_\infty = \frac{2}{7} v = \frac{2}{7} \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{2} \end{pmatrix}.$$